



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2009

Optimierung von 'sound quality' und Musikhören mit Cochlea-Implantaten

Lai, W K ; Dillier, N

Abstract: Cochlea-Implantate (CI) wurden in den letzten 20 Jahren bei vollständig und hochgradig tauben Patienten sehr erfolgreich zur Wiederherstellung des Sprachverständnisses eingesetzt, trotz Einschränkungen bezüglich Bandbreite, Klangdynamik sowie der zeitlichen und spektralen Auflösung im Vergleich zu Normalhörenden. Die meisten CI-Träger beurteilen jedoch die Klangqualität beim Musikhören als unbefriedigend und stark verbesserungswürdig. Unter der Annahme, dass eine Verbesserung der Klangqualität auch das Musikhören verbessert, wird hier überblicksmässig dargestellt, wie dieses Ziel mit der heutigen Technik zu erreichen ist und wo deren Grenzen liegen.

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: <https://doi.org/10.5167/uzh-46004>
Conference or Workshop Item

Originally published at:

Lai, W K; Dillier, N (2009). Optimierung von 'sound quality' und Musikhören mit Cochlea-Implantaten. In: 12. Jahrestagung der Deutschen Gesellschaft für Audiologie, Innsbruck, 11 March 2009 - 14 March 2009. DGA e.V., 1-4.

Optimierung von 'sound quality' und Musikhören mit Cochlea-Implantaten

Wai Kong Lai, Norbert Dillier

Labor für experimentelle Audiologie, ORL-Klinik, Universitätsspital Zürich, Schweiz

Schlüsselwörter: Musikwahrnehmung, Cochlea-Implantat

Einleitung

Cochlea-Implantate (CI) wurden in den letzten 20 Jahren bei vollständig und hochgradig tauben Patienten sehr erfolgreich zur Wiederherstellung des Sprachverständnisses eingesetzt, trotz Einschränkungen bezüglich Bandbreite, Klangdynamik sowie der zeitlichen und spektralen Auflösung im Vergleich zu Normalhörenden. Die meisten CI-Träger beurteilen jedoch die Klangqualität beim Musikhören als unbefriedigend und stark verbesserungswürdig.

Unter der Annahme, dass eine Verbesserung der Klangqualität auch das Musikhören verbessert, wird hier überblicksmässig dargestellt, wie dieses Ziel mit der heutigen Technik zu erreichen ist und wo deren Grenzen liegen.

Diskussion

Die Klangqualität wird vor allem durch die wahrgenommene Lautheit, Tonhöhe und zeitlichen Modulationen bestimmt. Die Lautheit wird insbesondere über die Eingangs(signal)- sowie Ausgangs(stimulations)-Dynamik kodiert, deren Kennlinien nichtlinear und patienten-spezifisch sind (Holden et al 2007, Spahr et al 2007).

Die Lautheitskodierung wird durch die akustische Eingangsdynamik sowie die elektrische Ausgangsdynamik beschränkt. Eine akustische Dynamik von 40-50dB genügt in der Regel, um Sprache gut zu repräsentieren. Musik hingegen benötigt für die naturgetreue Wiedergabe der leisesten bis lautesten Töne eine Dynamik von bis zu 90dB. Der CI-Sprachprozessor hat generell eine beschränkte Eingangsdynamik. Die elektrische Stimulation durch das CI beschränkt sich meist auf eine Ausgangsdynamik von 10 - 20 dB (die Differenz zwischen den T- und C-Werten der sogenannten Map im Prozessorprogramm). Die gesamte Komprimierung des Eingangs-Signals ist zwar akzeptabel für Sprache, jedoch nicht für Musik. Man kann teilweise diese Ausgangsdynamik noch vergrössern durch eine höhere Reizrate (die T-Werte werden niedriger, die C-Werte bleiben oft gleich), oder durch eine Desynchronisierung der neuronalen Aktivität mit einer sehr hohen Reizrate, welche als eine Art Conditioner wirkt (Hong & Rubinstein, 2003). Andererseits kann die Musikdynamik ohne allzu grosse Einbusse an Klangqualität verringert werden, was heutzutage sehr häufig gemacht wird, z.B. bei Radio- oder TV-Übertragungen oder auf kommerziellen Tonträger-Aufnahmen (z.B. CD).

Die Tonhöhe ist zum grossen Teil durch den Stimulationsort ("place pitch") kodiert, wobei die Anzahl sowie Position der Stimulations-Kanäle eine wichtige Rolle spielen. Gleichzeitige Stimulation benachbarter Elektroden (Busby & Plant 2005, Firszt et al 2007) kann zwar die Kanalzahl vermehren und eine feinere Frequenzzuordnung ermöglichen, verbessert jedoch wegen der elektrischen Feldausbreitung die Frequenzauflösung kaum (Busby et al 2008).

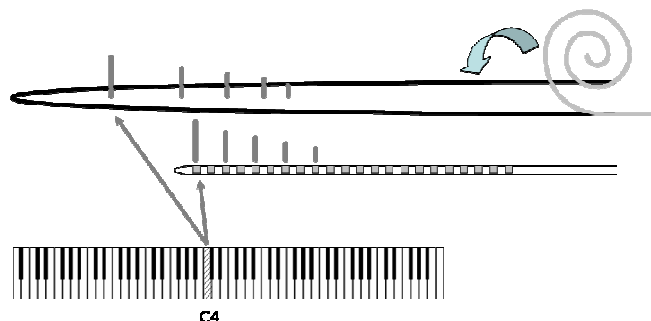


Abbildung 1: Dieses Beispiel zeigt die unterschiedlichen Stimulationsorte für akustische und elektrische Stimulation für den musikalischen Klang C4 mit Grundton (261 Hz) und dessen ersten 4 Obertönen. Das akustische Anregungsmuster wurde nach Greenwood (1990) berechnet. Das CI-Anregungsmuster entspricht dem Ausgangssignal für die ACE-Kodierungsstrategie (22 Kanäle, Frequenztafel 188-7938 Hz) und einer mittleren Insertion des Elektrodenträgers

Die Stimulationsorte auf der Basilar-Membran unterscheiden sich für akustische im Vergleich zu elektrischer Stimulation. Ein Beispiel für einen musikalischen Klang (C4) mit Grundton (261 Hz) und Obertönen ist in Abbildung 1 zu sehen. Für die Tonotopie der Basilar-Membran wurde die Greenwood-Formel (1990) eingesetzt, welche aber für elektrische Stimulation nicht korrekt ist und bezüglich der Lage der Ganglienzellen korrigiert werden muss (Briaire & Frijns 2006).

Damit ein Klang musikalisch und nicht geräuschartig wahrgenommen wird, ist die Erhaltung der harmonischen Struktur (Grundton und Obertöne) wichtig. Darüber hinaus sollten mittels einer geeigneten Frequenzzuteilung die musikalischen Töne gut voneinander trennbar sein, damit Melodien deutlich erkannt werden können. Eine "Halbtonzuteilung" (Omran et al 2009) ergab in einer experimentellen Studie zwar eine verbesserte Tonauflösung, jedoch eine subjektiv unangenehme Klangwahrnehmung.

Die MP3000-Kodierungsstrategie von Cochlear wurde in Anlehnung an das bekannte MP3-Komprimierungsverfahren entwickelt, ist aber an sich keine Musik-Kodierungsstrategie. Die Grundidee von MP3000 ist, die durch eine Maskierungsfunktion ermittelte redundante Stimulation wegzulassen. Dadurch kann ein vergleichbares Sprachverständnis mit einer geringeren Anzahl von Stimulationspulsen erreicht werden. Die „eingesparten Stimulationspulse“ können nun möglicherweise dazu verwendet werden, mehr spektrale Information zu übertragen. Für eine "*n*-of-*m*"-Kodierungsstrategie (SPEAK, ACE, MP3000) entspricht dieses Vorgehen der Erhöhung der Anzahl *n*. Die Ergebnisse einer Pilotstudie mit MP3000 und Musikwahrnehmung wurde an der DGA-Jahrestagung 2008 in Kiel vorgestellt (Lai & Dillier 2008).

Die Tonhöhen-Repräsentation durch den Stimulationsort ist jedoch grundsätzlich ungenügend, insbesondere wegen der elektrischen Felddausbreitung, welche zu "Channel-Interaction"-Effekten führt. Als erschwerend kommt hinzu, dass die Erregungs-Ausbreitung mit der Stimulationsstärke zunimmt und dadurch die Abbildung von Frequenzgruppen zu einzelnen Stimulationskanälen verschlechtert.

Die Wahrnehmung zeitlicher Modulationen hängt unter anderem von der Stimulationsrate ab, welche auch die Tonhöhe ("rate pitch") beeinflussen kann. Der "rate pitch" ist auf maximal 300 bis 1000 Hz beschränkt. Höhere Reizraten haben bisher nur wenig Verbesserung für die Sprachverständlichkeit (insbesondere im Störlärm) gebracht. Ob das auch für die Musikwahrnehmung zutrifft, bleibt eine offene Frage.

Die zeitliche Information kann in die Einhüllende (Envelope) sowie die zeitliche Feinstruktur (Temporal Fine Structure, TFS) aufgeteilt werden. Es ist wohl bekannt dass die Einhüllende mit dem CI gut wahrgenommen wird, und die für die Sprachdiskrimination ausreicht. TFS erfordert im Prinzip unterschiedliche Stimulationsrate auf den aktiven Elektroden. Für die zuverlässige Übertragung der zeitlichen Feinstruktur sind höhere, für einzelne aktive Elektroden unterschiedliche, Stimulationsraten sowie kanalsynchrone Reizmuster erforderlich. Ob die Übertragung der TFS-Information durch die "rate pitch"-Obergrenze eingeschränkt wird oder nicht, ist unklar. Erschwerend hinzu kommt eine mögliche zeitliche Verschmierung durch die Refraktär-Eigenschaften der elektrisch stimulierten Nervenfasern.

Einige Ansätze, die zeitliche Information mit CI besser zu übertragen, wurden bereits erarbeitet. Die "Fine Structure Processing-" (FSP-) Kodierungsstrategie (Riss et al. 2008) präsentiert die TFS Information in nur 2-3 (apikalen) Tieftönen-Kanälen und versucht dadurch, unterhalb der Obergrenze des "rate pitch" zu bleiben. Weitere Kodierungsstrategien (z.B. MEM, Vandali et al. 2005, und F0mod, Milczynski et al. 2009) versuchen, die F0-Modulation, einen weiteren Aspekt der zeitlichen Information, kanalsynchron und auch ausgeprägter zu präsentieren. Die Ergebnisse mit Musik mit sämtlichen Strategien sind bisher aber noch relativ beschränkt.

Was bei den erwähnten Ansätzen möglicherweise fehlt, ist eine gute Auflösung und Repräsentation der harmonischen und zeitlichen Struktur, wie sie bei normalem Gehör vorhanden sind. Vermutlich ein Hauptgrund für den Mangel an spektraler und temporaler Auflösung ist die elektrische Felddausbreitung, welche eine gute Ortsauflösung auch bei Erhöhung der Stimulationskanalzahl verhindert. Die Kodierung der zeitlichen Feinstruktur-Information muss noch besser untersucht werden, wird aber ebenfalls durch die elektrische Felddausbreitung eingeschränkt.

CI-Benutzer mit einem Restgehör können unter Umständen von einer kombinierten elektro-akustischen Stimulation (EAS) profitieren. Hier wird zusätzlich zur CI-Stimulation das Restgehör im Tieftönenbereich akustisch angeregt. Dabei kann die harmonische Grundstruktur des Musiksymbols sowie die TFS-Kodierung im Tieftönenbereich akustisch übermittelt und der Hochtonbereich durch das CI ergänzt werden. Eine entsprechende Anpassung und Aufteilung des Frequenzbereichs für die beiden Modalitäten erweist sich als sinnvoll. Das Ergebnis einer Musikstudie mit bimodaler Stimulation (Büchler et al. 2009) zeigt eine Verbesserung der subjektiven Klangqualität (Abbildung 2).

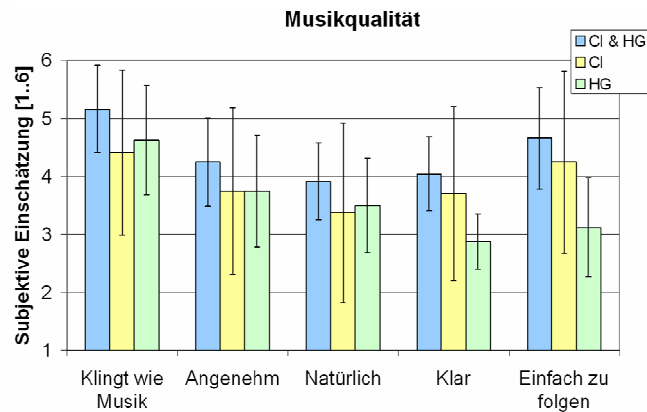


Abbildung 2: Subjektiv eingeschätzte Qualität der Musik, gemittelt über 6 Probanden. Wegen der grossen Streuung sind die Unterschiede zwischen den Bedingungen statistisch nicht signifikant. Ein Trend zeigt sich in den Kategorien "Klingt wie Musik", "Klar" und "Einfach zu folgen".

Zusammenfassung

Heutige CI-Systeme kodieren musikalische Klänge und Signale mit eingeschränkter Qualität. Die oben erwähnten Überlegungen weisen auf Entwicklungen hin, welche das Musikhören mit Cochlea-Implantaten künftig möglicherweise verbessern werden.

Literatur

- Briaire JJ, Frijns JHM. The consequences of neural degeneration regarding optimal cochlear implant position in scala tympani: A model approach. *Hear Res* 2006, 214(1-2): 17-27.
- Büchler M, Dillier N, Lai WK. Musikwahrnehmung bei bimodaler Stimulation. *Zeitschrift für Audiologie* 2009, 48(1): 16-25.
- Busby PA, Plant KL. Dual electrode stimulation using the Nucleus CI24RE cochlear implant: electrode impedance and pitch ranking studies. *Ear Hear* 2005, 26(5): 504-11.
- Busby PA, Battmer RD, Pesch J. Electrophysiological spread of excitation and pitch perception for dual and single electrodes using the Nucleus Freedom cochlear implant. *Ear Hear* 2008, 29(6): 853-64.
- Firszt JB, Koch DB, Downing M, Litvak L. Current steering creates additional pitch percepts in adult cochlear implant recipients. *Otol Neurotol* 2007, 28(5): 629-36.
- Holden LK, Skinner MW, Fourakis MS, Holden TA. Effect of increased IIDR in the Nucleus Freedom cochlear implant system. *J Am Acad Audiol* 2007, 18(9): 777-93.
- Hong RS, Rubinstein JT. High-rate conditioning pulse trains in cochlear implants: Dynamic range measures with sinusoidal stimuli. *J Acoust Soc Am* 2003, 114(6): 3327-42.
- Lai WK, Dillier N. Investigating the MP3000 coding strategy for music perception. Tagungs-CD DGA 11. Jahrestagung, Kiel 05.-08. März 2008.
- Milczynski M, Wouters J, van Wieringen A. Improved F0 Coding in Cochlear Implant Signal Processing. *J Acoust Soc Am* 2009, 125(4): 2260-2271.
- Omran S, Lai WK, Büchler M, Dillier N. Melody contour identification and instrument recognition using semitone mapping in Nucleus Cochlea Implant recipients. Präsentation DGA 12. Jahrestagung, Innsbruck 11.-14. März 2009.
- Riss D, Arnolder C, Baumgartner WD, Kaider A, Hanzavi JS. A new fine structure speech coding strategy: speech perception at a reduced number of channels. *Otol Neurotol* 2008, 29(6): 784-8.
- Spahr AJ, Doman MF, Loisele LH. Performance of patients using different cochlear implant systems: effects of input dynamic range. *Ear Hear* 2007, 28(2): 260-75.
- Vandali AE, Sucher C, Tsang DJ, McKay CM, Chew JW, McDermott HJ. Pitch ranking ability of cochlear implant recipients: a comparison of sound-processing strategies. *J Acoust Soc Am* 2005, 117(5): 3126-38.

Acknowledgement

Diese Arbeit wurde unterstützt durch den Schweizer Nationalfonds zur Förderung der Wissenschaftlichen Forschung (Grant Nr. 320000-110043), sowie ein Forschungs-Stipendium der Firma Cochlear Cochlear AG, Basel.